Spontaneous Mutation at a 5-Methylcytosine Hotspot Is Prevented by Very Short Patch (VSP) Mismatch Repair

Margaret Lieb

Department of Microbiology, University of Southern California School of Medicine, Los Angeles, California 90033

Manuscript received July 9, 1990
Accepted for publication January 8, 1991

ABSTRACT

In many strains of *Escherichia coli,* the product of gene *dcm* methylates the internal cytosines in the sequence 5′CC(A or T)GG. Spontaneous deamination of 5-methylcytosine produces thymine which, if not corrected, can result in a transition mutation. 5-Methylcytosines in the *lacI* gene are hotspots for spontaneous C to T mutations. *dcm* is linked to *vsr,* a gene required for very short patch (VSP) repair. VSP repair corrects T·G mispairs in the following contexts: $\frac{CTAGG}{GGTCC}$, $\frac{CTTGG}{GGACC}$, $\frac{TAGG}{GTCC}$ and $\frac{CTAG}{GGTC}$. I have investigated the relationships between cytosine methylation, mutation, and VSP repair. Spontaneous mutations in the repressor (*cI*) gene of lambda prophage were isolated in wild-type and mutant lysogens. A hotspot for spontaneous mutation that corresponds with a 5-methylcytosine was observed in wild-type lysogens but was not present in bacteria lacking both methylase and VSP repair activity. Introduction of a plasmid containing dcm^+ and vsr^+ restored the mutation hotspot. If the added plasmid carried only dcm^+, the frequency of spontaneous mutations at the 5-methylcytosine was over 10-fold higher than in Dcm^+Vsr^+ lysogens. The addition of *vsr* on a plasmid to a wild-type lysogen resulted in a 4-fold reduction in mutation at the hotspot. These findings support the previously untested hypothesis that VSP repair prevents mutations resulting from deamination of 5-methylcytosine.

IN the *lacI* gene of *Escherichia coli,* three sites where spontaneous C to T mutations occur frequently coincide with cytosines subject to methylation (COULONDRE *et al.* 1978). Inasmuch as the hotspots were not observed in *E. coli* B, a strain lacking cytosine methylase activity, deamination of the 5-methylcytosines in *E. coli* K12 is presumed to be responsible for the mutation hotspots (COULANDRE *et al.* 1978). *E. coli* K12 also contains a specific mismatch repair function that recognizes the TG pairs that would result from deamination of 5-methylcytosine, and replaces T with C. The repair system was called VSP (very short patch) repair because correction of the mismatch requires the removal of less than 5 residues (LIEB 1983; LIEB, ALLEN and READ 1986). VSP repair accounts for "high negative interference" between closely linked markers when one of the markers is a hotspot mutation, because repair-prone T·G mispairs arise in DNA heteroduplexes that accompany recombination (LIEB 1983). T·G mispairs in lambda heteroduplexes prepared *in vitro* are also subject to VSP repair after transfection of the DNA into wild-type K12 bacteria (LIEB 1987; JONES, WAGNER and RADMAN 1987; MESELSON 1988). Recently, a gene required for VSP repair has been identified (BHAGWAT, SOHAIL and LIEB 1988a). This gene, *vsr,* lies immediately downstream from *dcm* in the same operon (SOHAIL *et al.* 1990) Consequently, certain mutations, such as *dcm6,* lead to the inactivation of both methylation and repair activities in the bacterium (LIEB 1987; JONES, WAGNER and RADMAN 1987; ZELL and FRITZ 1987). However, VSP repair does not require the product of *dcm* (SOHAIL *et al.* 1990).

VSP repair is assumed to have an antimutagenic function in bacteria that contain 5-methylcytosine. The existence of mutation hotspots at 5-methylcytosine in *lacI* in a dcm^+vsr^+ strain, which is expected to be proficient in VSP repair, suggests three hypotheses: (1) VSP repair does not prevent spontaneous mutations. In that case, one would expect the hotspots to be unaffected by the presence or absence of *vsr*; (2) potential mutations occur at a very high rate and VSP does not correct them all. Additional copies of gene *vsr* might then reduce mutation at the hotspot in wild-type bacteria; and (3) mutation hotspots observed in wild-type bacteria are not caused by deamination of 5-methylcytosine, resulting in T·G mispairs, but by some other property of the modified base, such as mispairing with A during replication. This model predicts that the hotspot would persist even if VSP repair were increased by adding additional copies of *vsr* to wild-type bacteria.

MATERIALS AND METHODS

Bacteria, bacteriophage, and plasmids: Strain GM3803 ($sup0$ dcm^+vsr^+) is a derivative of GM30 (MARINUS 1973). GM2142 is an isogenic strain carrying mutation *dcm6*, which results in a loss of cytosine methylase activity (MARINUS and

MORRIS 1973) and in a large reduction in VSP repair activity (LIEB 1987, JONES, WAGNER and RADMAN 1987). Strain C600, which carries *supE* (BACHMAN 1972), was used to prepare phage stocks. Bacteriophage lambda mutants and mapping methods were described previously (LIEB 1981). Mutation *ind6* in gene *cI* prevents repressor inactivation by the SOS response (reviewed by WALKER 1984). Plasmids pDCM1 and pDCM23 are derivatives of pBR322 containing wild type *E. coli* genes *dcm* and *vsr*, and different amounts of adjacent bacterial DNA (BHAGWAT, SOHAIL and LIEB 1988b). Plasmid pDCM21 contains *dcm*, with the deletion of the distal two-thirds of *vsr* (BHAGWAT, SOHAIL and LIEB 1988b). Plasmid pvsr20 is pET3a (SAMBROOK, FRITSCH and MANIATIS 1989) containing vsr^+ and no additional *E. coli* genes (A. BHAGWAT, personal communication). Plasmids were introduced into lysogens as described by MANIATIS FRITSCH and SAMBROOK (1982).

Preparation of lysogens: Bacteriophage lambda *cIind6* phage was spotted on bacterial lawns. Lysogens were isolated by streaking from the turbid center of the phage spot. Clones of lysogenic bacteria were grown in liquid medium; supernatent medium from each culture was tested for the presence of phage. Spontaneous derepression and subsequent excision of Ind^- prophage is very rare, so that most phage produced by lysogens containing a single integrated prophage per genome have sustained a mutation that inactivates the repressor. Mutants unable to adsorb lambda were selected using lambda i^{434}: resistance to phage adsorption is required to prevent the loss of free (mutant) phage particles arising in lysogenic cultures.

Culture conditions and mutant selection: Single colonies were used to prepare stock cultures of lysogens. The bacteria were then diluted in broth, and 0.2-ml aliquots (containing about 10^4 bacteria) were distributed into small tubes. A total of at least 20 sets of 20–50 tubes were generally inoculated with each lysogen. The tubes were incubated until turbidity was visible; some sets of tubes were then agitated on a rotary shaker at 37–41° for several hours until the bacterial concentration reached about 10^9 per ml (2×10^8 per tube). The cultures were treated with chloroform and plated on SupO bacteria (LIEB 1981).

Identifying amber and IS mutants: The number of plaques produced by individual 0.2 ml minicultures ranged from zero to several thousand. Up to 20 plaques from each plate were screened for amber mutations as follows. A sterile toothpick was used to stab a single plaque, and then to stab in succession plates spread with SupO, SupE and SupF bacteria. Unless ambers with different patterns of suppression were found, only one amber mutant was isolated from any miniculture. Amber mutants were plaque-purified, stocks were grown, and the mutation was identified by crossing with phages bearing previously mapped *cI* mutations (Table 1). For each lysogen studied, 100 or more independent spontaneous *cI* mutants picked at random were tested for the presence of IS elements by exposure to 0.01 M $Na_4P_2O_4$, which rapidly inactivates phage particles containing genomes with large insertions (LIEB 1981).

Calculations: For each set of minicultures, the fraction of tubes containing no phage (P_0) was used to calculate **a**, the average number of mutations per tube, using Poisson's equation: $P_0 = e^{-a}$. Multiplying **a** by the number of tubes in the set gives the total number of mutations. In sets in which a particular mutation (*am6*) was found in 20% or more of the tubes, a correction was made for the occurrence of more than one such mutation per miniculture.

Calculations of true mutation rates is problematic, since there is a lag of several bacterial divisions between the mutation event that results in an inactive repressor, and the appearance of bacteria lacking sufficient repressor to prevent prophage induction. Mutation rates were calculated using the average number of mutations per culture as described by LURIA and DELBRÜCK (1943).

RESULTS

A mutation hotspot is present at a 5-methylcytosine in gene *cI*: Spontaneous lambda *cI* mutants were isolated from a lysogen that is dcm^+ (GM3803), and from an isogenic *dcm6* lysogen (GM2142) which is defective in cytosine methylation and VSP repair. Among 51 amber mutants isolated in the dcm^+vsr^+ strain, 26 were *identical to am6*, which is the result of a C to T transition in the sequence 5'CmeCAGG (LIEB, ALLEN and READ 1986). The *am6* mutations accounted for about 2% of all mutations tested (Table 2, line 1). In the *dcm6* strain, the fraction of mutations that occurred at *am6* was less than 0.2% (Table 2, line 2). When *am6* mutations are excluded from the calculation, the fraction of *cI* mutations that were ambers was very similar in dcm^+ and *dcm6* lysogens (Table 2). The association of a mutation hotspot with 5-methylcytosine is in accord with the observations of COULONDRE *et al.* (1978).

To confirm the association between methylation and the mutation hotspot, a plasmid containing dcm^+ was inserted into GM2142. As shown on line 3 of Table 2, the hotspot at *am6* reappeared.

Mutation at 5-methylcytosine increases in the absence of VSP repair: Because a *vsr* mutant that retains cytosine methylase activity was not available, this phenotype was constructed by transferring pDCM21, which carries dcm^+ but not vsr^+, into a *dcm6* lysogen. The resistance to *Eco*RII digestion of DNA isolated from the transformant confirmed its Dcm^+ phenotype (data not shown). Almost all of the *cI* amber mutants arising in this strain were *am6*, and over 22% of all *cI* mutations were at the hotspot (Table 2, line 4). This is 12 times the frequency of *am6* among spontaneous mutations arising in the wild-type lysogen. In a wild-type lysogen carrying a dcm^+vsr^+ plasmid, only 1.3% of spontaneous mutations were at the hotspot (Table 2, line 5). Thus, the large increase in mutation at the hotspot is not due to the presence of dcm^+ on a plasmid, but is correlated with the absence of gene *vsr*.

Additional copies of vsr^+ reduce mutation at 5-methylcytosine: Although it has been shown that 5-methylcytosine is much more mutagenic in the absence of *vsr* function than in its presence, the presence of mutation hotspots in wild-type (dcm^+vsr^+) bacteria requires an explanation. In lysogens containing plasmids carrying both dcm^+ and vsr^+ (Table 2, lines 3 and 5), the frequency of *am6* mutations was lower than in wild-type bacteria, suggesting that an increase in Vsr product reduced mutation at the hot spot. In plasmid pvsr20, vsr^+ is expressed from the strong T7

TABLE 1

Spontaneous *cI* amber mutations detected in *E. coli* lysogens

amber mutation sites		Number of occurrences in each host					
Site	Sequence (5′ → 3′)	1 GM3803	2 GM2142	3 GM2142 (pDCM1)	4 GM2142 (pDCM21)	5 GM3803 (pDCM23)	6 GM3803 (pvsr20)
6	T C C **C** A G G	26	5	6	122	22	7
289	A G T **G** G C C	5	5	0	0	0	4
B36	T C T **G** G C T	4	2	0	0	2	1
212	G T A **T** G A G	3	2	0	0	0	0
10	G T T **C** A G G	3	4	0	1	0	0
9	G A G **C** A G G	2	6	1	0	0	2
L8	G A T **G** A G T	2	2	0	0	0	0
1	G A T **G** G G T	2	4	1	3	4	3
15	G G T **C** A G G	0	0	0	0	0	0
282	G G G **C** A G T	1	4	0	0	2	0
505	C C A **C** A G T	1	3	0	0	2	1
500	C C T **G** A G C	1	1	0	0	0	0
273	C T A **C** G A G	1	1	0	0	0	0
J7	A T G **C** A G C	0	4	0	0	0	0
Q18	C T T **G** A G G	0	2	0	0	0	0
14	A A T **G** A G A	0	2	0	0	0	0
110	A G T **G** A G T	0	2	0	1	0	0
325	T A T **G** A G T	0	1	0	0	0	0
499	G T A **C** C C T	0	1	0	0	0	0
p21R4	G A A **G** A G A	0	0	1	1	0	0
C58	G A G **C** A G C	0	0	0	0	1	1
12	C T T **G** A G G	0	0	0	0	0	1
288	A G T **C** A G T	0	0	1	0	0	0
Total ambers		51	51	10	112	33	20
Total no. of mutations screened		1408	2917	608	542	1663	1629

TABLE 2

Effect of *dcm* and *vsr* on the frequency of *cI* mutations

	Lysogen		Plasmid		Percent *am6*	Percent other ambers	Percent IS	"Mutation rate" ($\times 10^{-8}$)
	Strain	Genotype	Name	Genotype				
1	GM3803	*dcm*$^+$*vsr*$^+$			1.85	1.78	48	1.3 ± 0.1
2	GM2142	*dcm6*			0.17	1.58	57	1.1 ± 0.4
3	GM2142	*dcm6*	pDCM1	*dcm*$^+$*vsr*$^+$	0.99	0.82	ND	ND
4	GM2142	*dcm6*	pDCM21	*dcm*$^+$	22.5	1.10	35	1.3 ± 0.3
5	GM3803	*dcm*$^+$*vsr*$^+$	pDCM23	*dcm*$^+$*vsr*$^+$	1.3	0.66	51	ND
6	GM3803	*dcm*$^+$*vsr*$^+$	*pvsr20*	*vsr*$^+$	0.43	0.80	56	0.8 ± 0.2

promoter. Mismatch repair of *am6* during phage crosses increased about 2-fold in bacterial hosts containing *pvsr20* (M. LIEB, unpublished observations). Addition of *pvsr20* to wild-type lysogens resulted in a 4-fold drop in the frequency of *am6* (Table 2, line 6). Thus, we may conclude that most of the mutations at 5-methylcytosine in wild-type bacteria can be prevented by VSP repair, and thus are the result of cytosine deamination.

DISCUSSION

In the repressor gene of lambda, as in the bacterial *lacI* gene, 5-methylcytosine is a hot spot for C to T transitions. When cytosine methylation was ablated by a *dcm* mutation, there was a 10-fold reduction in the frequency of mutation at the hotspot. A similar reduction in the relative number of spontaneous mutations at hotspots was observed by COULONDRE *et al.* (1978) when they compared spontaneous mutation in *E. coli* strains K12(Dcm$^+$) and B (Dcm$^-$).

Mutation at 5-methylcytosine was 10-fold more frequent in a strain lacking Vsr function than in a Dcm$^+$Vsr$^+$ lysogen. Although VSP repair has been assumed to have a role in the prevention of spontaneous mutation (JONES, WAGNER and RADMAN 1987; ZELL and FRITZ 1987), this is the first direct demonstration of such a function. Since VSP specifically

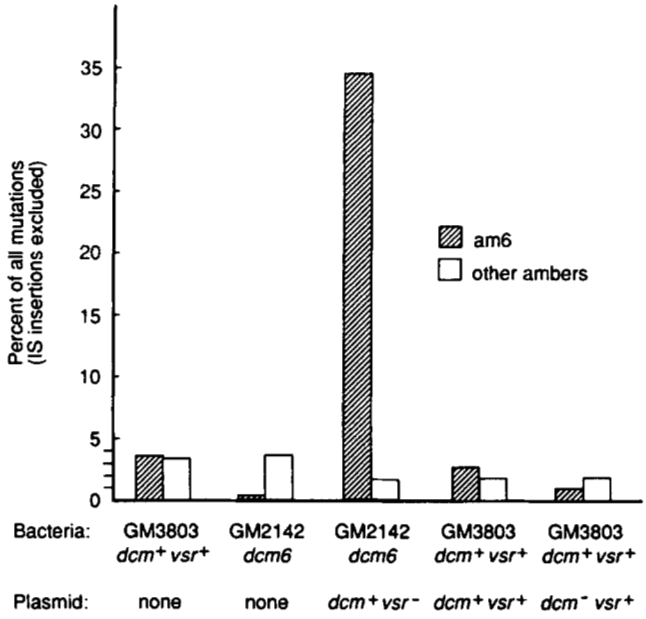

FIGURE 1.—Effect of VSP repair on the frequency of spontaneous amber mutations in λ gene *cI*. Results displayed are for lysogens 1, 2, 4, 5 and 6 listed in Table 1.

corrects T·G mispairs, this also shows that deamination of 5-methylcytosine is a mutagenic event *in vivo*. Obviously, VSP repair does not prevent all mutations at 5-methylcytosine, because hotspots are found in Dcm^+Vsr^+ bacteria. The apparent inefficiency of repair is attributable to insufficient Vsr product because the *dcm6* hotspot was greatly reduced in bacteria in which a plasmid copy of vsr^+ was expressed efficiently. Thus, we can eliminate hypothesis 1, noninvolvement of Vsr in mutation prevention, and we have strong support for hypothesis 2, a deficiency of Vsr in wild-type bacteria. The fact that overexpression of *vsr* did not reduce mutation at 5-methylcytosine to the level observed in dcm^- bacteria suggests that events other than cytosine deamination might be responsible for some of the mutagenicity of the modified base (hypothesis 3). However, such events could account for only a small fraction of hotspot mutations.

In order to determine true mutation rates in the system used in this study, one would have to determine the number of generations that intervene between the occurrence of a mutation in a prophage of a lysogen and the production of phage progeny by a descendent of that lysogen. This "phenotypic lag" is probably over 4 bacterial generations, and may be quite variable. The "mutation rates" presented in Table 2 were calculated using the number of bacteria per culture at the time they were plated to detect free phage, and not the number at the time the mutation occurred. However, the calculations clearly show that there are no significant differences in gene *cI* mutation rates in the various lysogens.

Because the sequence of gene *cI* makes it prone to two special types of mutation, increased mutation at a hotspot is not reflected in the overall rate of mutation. The gene is a preferred target for the transposition of IS elements, and about 50% of all cI mutations detected were IS insertions (Table 2, see also LIEB 1983). If we eliminate IS mutations from consideration, we see that the spectrum of point mutations is noticeably affected by the increased number of mutations at the 5-methylcytosine hot spot (Figure 1). In the dcm^+vsr^- lysogen, almost 35% of non-IS mutations were at the hotspot. The increase in hotspot mutations is reflected in a decrease in the fraction of the total mutations that were IS insertions (Table 2, line 4). Since about 60% of non-IS mutations are frameshifts occurring in runs of T·A base pairs (LIEB 1983), mutations at 5-methylcytosine are a very large fraction of the missense and nonsense mutations that occur in the absence of VSP repair.

An increase in mutation at 5-methylcytosines in *lacI* after conjugal transfer of the gene was observed by KUNZ and GLICKMAN (1983). Infidelity of replication of a damaged template was suggested as a possible explanation. Loss of VSP repair function during conjugation would also account for preferential mutation at the hotspots. In addition, deamination of 5-methylcytosine has been reported to occur more frequently when DNA is denatured (LINDAHL 1979).

Cytosine methylation occurs in most eukaryotes, usually in CpG sequences. 5-methylcytosine apparently has important roles in gene expression and development (reviewed by CEDAR 1988), and it is a significant source of mutations that lead to inherited diseases (COOPER and YOUSSOUFIAN 1988). A mismatch repair system to prevent the loss of 5-methylcytosine by deamination would be advantageous in higher organisms. An activity that corrects T·G mispairs to C·G by replacing a single base has been demonstrated in extracts of human cells (BROWN and JIRICNY 1987; JIRICNY *et al.* 1988; WIEBAUER and JIRICNY 1989). However, the repair is not limited to mispairs in a particular context. As is the case with VSP repair in *E. coli*, the mammalian repair system may not succeed in correcting all mispairs caused by events at 5-methylcytosine. It will be interesting to determine whether there is any homology between prokaryotic and eukaryotic VSP repair systems.

This investigation was supported by grants DMB 85–16062 from the National Science Foundation and GM 25034 from the National Institutes of Health. I thank M. MARINUS for supplying bacterial strains and A. BHAGWAT for contributing plasmids and valuable discussions.

LITERATURE CITED

BACHMAN, B. J., 1972 Pedigrees of some mutant strains of *Escherichia coli* K-12. Bacteriol. Rev. **36:** 525–557.

BHAGWAT, A. S., A. SOHAIL and M. LIEB, 1988a A new gene involved in mismatch correction in *Escherichia coli*. Gene **74:** 155–156.

BHAGWAT, A. S., A. SOHAIL and M. LIEB, 1988b Methylation abilities of *Eco*RII and Dcm methylases are not sufficient for their involvement in very short patch repair, pp. 173–178 in *Mechanisms and Consequences of DNA Damage Processing*. Alan R. Liss, New York.

BROWN, T. C., and J. JIRICNY, 1987 A specific mismatch repair event protects mammalian cells from loss of 5-methylcytosines. Cell **50:** 945–950.

CEDAR, H., 1988 DNA methylation and cell activity. Cell **53:** 3–4.

COOPER, D. N., and H. YOUSSOUFIAN, 1988 The CpG dinucleotide and human genetic disease. Hum. Genet. **78:** 151–155.

COULONDRE, C., J. H. MILLER, P. J. FARABAUGH and W. GILBERT, 1978 Molecular basis of base substitution hotspots in *Escherichia coli*. Nature **274:** 775–780.

JIRICNY, J., M. HUGHES, N. CORMON and B. B. RUDKIN, 1988 A human 200-kDa protein binds selectively to DNA fragments containing G·T mismatches. Proc. Natl. Acad. Sci. USA **85:** 8860–8864.

JONES, M., R. WAGNER and M. RADMAN, 1987 Mismatch repair of deaminated 5-methylcytosine. J. Mol. Biol. **193:** 155–159.

KUNZ, B. A., and B. W. GLICKMAN, 1983 The infidelity of conjugal DNA transfer in *Escherichia coli*. Genetics **105:** 489–500.

LIEB, M., 1981 A fine structure map of spontaneous and induced mutations in the lambda repressor gene, including insertions of IS elements. Mol. Gen. Genet. **184:** 364–371.

LIEB, M., 1983 Specific mismatch correction in bacteriophage lambda crosses by very short patch repair. Mol. Gen. Genet. **191:** 118–125.

LIEB, M., 1987 Bacterial genes *mutL, mutS* and *dcm* participate in repair of mismatches at 5-methylcytosine sites. J. Bacteriol. **169:** 5241–5246.

LIEB, M., E. ALLEN and D. READ, 1986 Very short patch mismatch repair in phage lambda: repair sites and lengths of repair tracts. Genetics **114:** 1041–1060.

LINDAHL, T., 1979 DNA glycosylases, endonucleases for apurinie/apyrimidinic sites, and base excision-repair. Prog. Nucleic Acid Res. Mol. Biol. **20:** 135–192.

LURIA, S. E., and M. DELBRÜCK, 1943 Mutation of bacteria from virus sensitivity to virus resistance. Genetics **28:** 491–511.

MANIATIS, T., E. F. FRITSCH and J. SAMBROCK, 1982 *Molecular Cloning, A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

MARINUS, M. G., 1973 Location of DNA methylation genes on the *Escherichia coli* K-12 genetic map. Mol. Gen. Genet. **127:** 45–55.

MARINUS, M. G., and N. R. MORRIS, 1973 Isolation of DNA methylase mutants from *Escherichia coli* K-12. J. Bacteriol. **114:** 1143–1150.

MESELSON, M., 1988 Methyl-directed repair of DNA mismatches, pp 91–113 in *The Recombination of Genetic Material*, edited by K. B. LOW. Academic Press, New York.

SAMBROOK, J., E. F. FRITSCH and M. MANIATIS, 1989 *Molecular Cloning*, Ed. 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

SOHAIL, A., M. LIEB, M. DAR and A. S. BHAGWAT, 1990 A gene required for very short patch repair in *Escherichia coli* is adjacent to the DNA cytosine methylase gene. J. Bacteriol. **172:** 4214–4221.

WALKER, G. C., 1984 Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. Microbiol. Rev. **48:** 60–93.

WIEBAUER, K., and J. JIRICNY, 1989 *In vitro* correction of G·T mispairs to G·C pairs in nuclear extracts from human cells. *Nature* **339:** 234–236.

ZELL, R., and H. FRITZ, 1987 DNA mismatch-repair in *Escherichia coli* counteracting the hydrolytic deamination of 5-methyl-cytosine residues. EMBO J. **6:** 1809–1815.

Communicating editor: P. J. PUKKILA